

DRAFT

**MIML Development
Version 0.3 – 25 June 2003**

**Maritime Information Markup Language
Guidelines for Development and Maintenance**

DRAFT

TABLE OF CONTENTS

1. PURPOSE.....	3
2. SCOPE	3
3. BACKGROUND	3
4. WIN	3
5. XML	5
6. MIML MANAGEMENT APPROACH.....	5
6.1 WIN/MIML Configuration Control Board	5
<i>6.1.1 WIN/MIML CCB Duties.....</i>	<i>6</i>
6.2 MIML Models	6
<i>6.2.1 Minimal Model/Schema Requirements (Metadata)</i>	<i>6</i>
6.3 MIML Model Registry	6
<i>6.3.1 Searching the MIML Registry.....</i>	<i>8</i>
6.4 Maintenance of the MIML	8
<i>6.4.1 Identify MIML Change.....</i>	<i>9</i>
<i>6.4.2 Review and Assign</i>	<i>9</i>
<i>6.4.3 Develop MIML Change.....</i>	<i>10</i>
6.5 Update MIML Documentation	10
6.6 MIML Assumptions.....	10
6.7 Current MIML.....	11
6.8 Future MIML	11
6.9 Outstanding MIML Issues	12
6.10 WIN/MIML CCB Recommendations	12
7. MIML TECHNICAL APPROACH.....	12
7.1 Overall structure of technical solution.....	12
<i>7.1.1 MIML Functions</i>	<i>13</i>
<i>7.1.2 User Levels of Authority</i>	<i>13</i>
7.2 MIML Knowledge Models	13
<i>7.2.1 Basic Schemas.....</i>	<i>13</i>
<i>7.2.2 Derived Schemas.....</i>	<i>13</i>
<i>7.2.3 Schema Management</i>	<i>14</i>
7.3 Dictionary/Registry of Knowledge Models.....	14
<i>7.3.1 MIML Model Registry Submission Process.....</i>	<i>14</i>
<i>7.3.2 Preparation of submission package for the MIML registry</i>	<i>14</i>
<i>7.3.3 MIML Model Validation Rules</i>	<i>15</i>
<i>7.3.4 Basic Structure of the Registry for the MIML Models.....</i>	<i>15</i>
<i>7.3.5 Version Control.....</i>	<i>16</i>
<i>7.3.6 Model or Schema Derivation.....</i>	<i>16</i>
7.4 Model Redundancy	16
7.5 Model Integration/Reconciliation.....	16
8. MIML DEFINITIONS	17
9. REFERENCES.....	17

1. PURPOSE

This document is intended to provide a strategy and guidance on the use of Extensible Markup Language (XML) within the Waterways Information Network (WIN) that is being developed by the Coast Guard Research and Development Center (RDC).

2. SCOPE

This document applies to all government and commercial organizations that are engaged in developing, acquiring, or maintaining systems that contribute to the WIN. The goal of WIN is to have XML used to describe the format of data transfer mechanisms and structure of data available on the WIN. These descriptions will be known as MIML. This document addresses MIML implementation as it applies to automated systems, applications, data exchanges, databases and information presentations within and across all operations represented within the WIN. Customers that use data from WIN will know what kind of data is available. They will be looking for original and recent data for applications they use. The WIN transfer mechanism will not provide descriptions of “how” data would be used (i.e. The data will not be self-describing).

3. BACKGROUND

The Coast Guard (CG) recognized the need to address maritime information issues that involve other agencies, industry, and our own concerns. Three Commandant initiatives resulted:

- Memorandum of Agreement (MOA) – NOS/NOAA and USCG – to cooperate on accelerating the production of S-57 Electronic Charts and other related matters.
- Electronic Chart Guidance Team (ECGT) – Chartered to address and develop solutions to various problems associated with electronic charts and updating.
- Marine Information Exchange Business Solutions Team (MIE-BST)

By the MOA, NOAA is a key stakeholder on the ECGT. Though a major purpose of the ECGT was to ensure internal CG operability with electronic chart progress, other problems were noted. NOAA raised an immediate issue concerning Notice to Mariners information generated by the CG. *The Guidance Team noted “the entire data flow is antiquated, requiring human intervention at several steps, introducing an unacceptable error rate,”* and though “several ideas were discussed as to how to fix the problem,” little actual progress has been made.

The MIE-BST was chartered to “identify the technologies, processes, policies and resources necessary to maximize the Public to Coast Guard (P2CG) and Coast Guard to Public (CG2P) interface and infrastructure at *Coast Guard Groups, Activities and Marine Safety Offices.*”

4. WIN

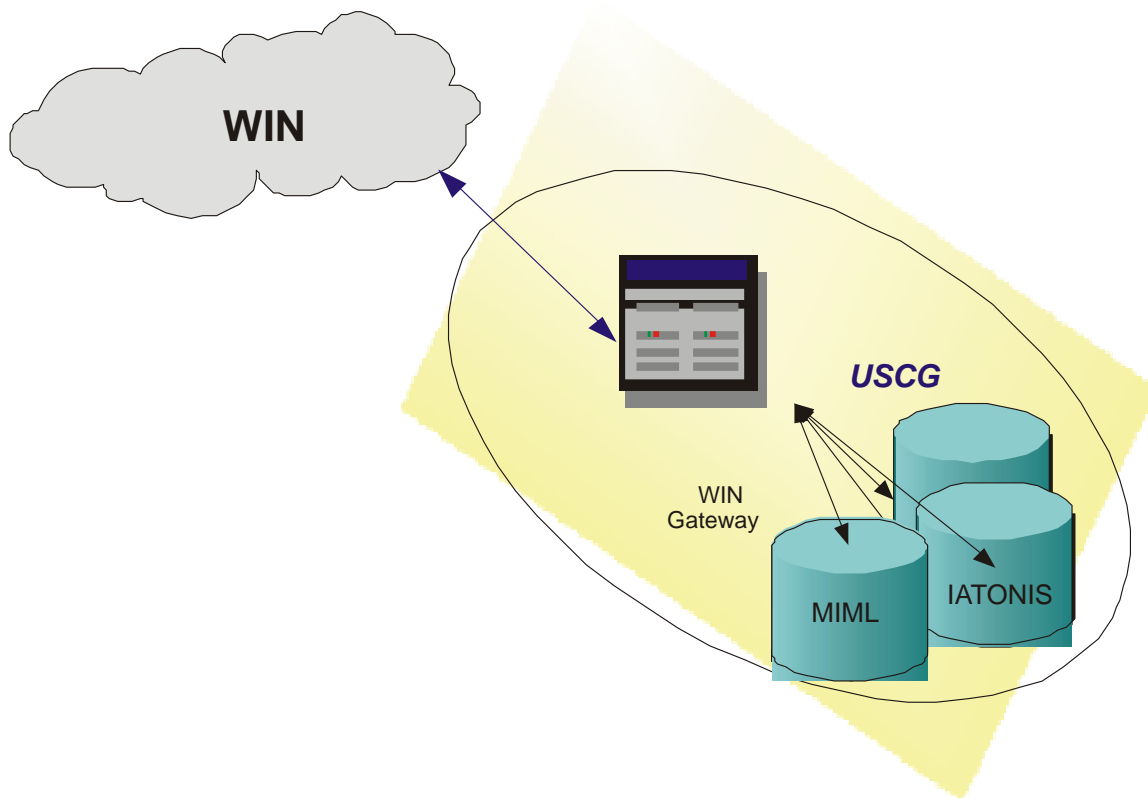
This is the prototype system that the Coast Guard RDC is developing to focus on Coast Guard problems in maritime information. This prototype is an Internet-based, distributed

DRAFT

network, open architecture system that has the potential to best meet foreseeable Coast Guard needs. This solution provides the means for information users to directly access data from information providers; to the extent the provider makes the information available. The system will allow direct processing of information and will provide an open architecture for application growth. With WIN, the Coast Guard has an opportunity to improve information sharing that will significantly reduce costs of situational awareness and decision support tools and enable even more, new capabilities. This prototype may include information transfers to other agencies. Once benefits to the Coast Guard are achieved, a logical extension of the benefits would reach throughout government agencies and the MTS community as a whole.

- The MIML Knowledge Base that contains all the MIML [models](#), and basic [schema](#) is located behind the WIN Gateway and is available to any user of the WIN.
- The WIN Peer to Peer protocols and frameworks that use XML to markup message header information necessary for binding, reliable messaging and security will also be included in the MIML. These along with the basic schema become the MIML vocabulary specification.
- The documents that use the MIML will be maintained on the WIN in the form of XML pages. They will contain the actual business information communicated.

Figure 1



5. XML

The Extensible Markup Language originated within the World Wide Web Consortium (W3C) as a semi-structured data exchange format that included both data and a description of the data's structure in a single package. A number of W3C technical specifications have been developed that define XML. Reference 9 is the core specification that provides syntax rules for using XML for a variety of data exchange, presentation, storage, protocol development and other purposes. References 7 and 8 provide XML-based mechanisms for defining specified formats for XML data exchanges. A listing of all W3C Technical Specifications can be found at <http://www.w3.org>.

The Coast Guard RDC will be using the guidelines provided by the U.S. Federal CIO Council XML Working Group in the Draft Federal XML Developer's Guide: Version 0.2, Reference 4, which is an approved adaptation of the consensus draft of the Department of the Navy (DON) XML Developer's Guide (version 1.1) of November 7, 2001, Reference 6. Most of the definitions used in this document can be found in their glossaries

6. MIML MANAGEMENT APPROACH

6.1 WIN/MIML Configuration Control Board

A variety of authoritative sources of information have a direct bearing on the structure and content of MIML. A change in any one of these sources will likely cause a change to MIML. In addition, a WIN Partner may initiate a change request to MIML. A WIN/MIML Configuration Control Board (CCB) will be established. It will be made up of WIN Partners and WIN/MIML Administrators. It will have representatives from all subgroups and the users (application designers/programmers). The WIN/MIML CCB will be at the center of the MIML process. The Board will be supported by the WIN Administrator, who will act as executive secretary for the Board, and the MIML Administrator.

The Board will identify needs and initiate guidance for enhancements or corrections to the MIML, and will approve any changes. It has overall charge of all the [models/schemas](#), and maintenance groups for specific models/schemas. The WIN/MIML CCB would vet changes in individual models/schemas for effects on others. Only the maintenance group assigned may do any update to that specific model/schema, and must be approved by the WIN/MIML CCB before acceptance. The WIN/MIML CCB will review WIN changes for impact, assign them for action, and monitor progress. When these changes have been implemented in the contributing system, the appropriate changes to the MIML will be made and posted on the WIN. The WIN Partners will then be notified.

6.1.1 WIN/MIML CCB Duties

1. Develop MIML Criteria
2. Recognize problems with [models](#) (Consumer Report)
3. Test models against MIML Criteria
4. Provide evaluations of models
5. Review [owner](#) evaluations of models
6. Review user evaluations of models
7. Recommend models (i.e. geospatial)

6.2 MIML Models

A MIML [Model](#) must meet the minimum requirements. There is a range of sophistication allowed in the MIML models. A minimal model will be a document with “tags” in front of the real data. A second level of model will be the “tagged” document with an associated document that describes the tags and gives examples of them. A third level of sophistication in MIML would be an Entity-Relationship (E/R) model, a model developed with the Unified Modeling Language (UML) or any other kind of model developed with a modeling tool such as Protégé. See Figure 2 for examples.

6.2.1 Minimal Model/Schema Requirements (Metadata)

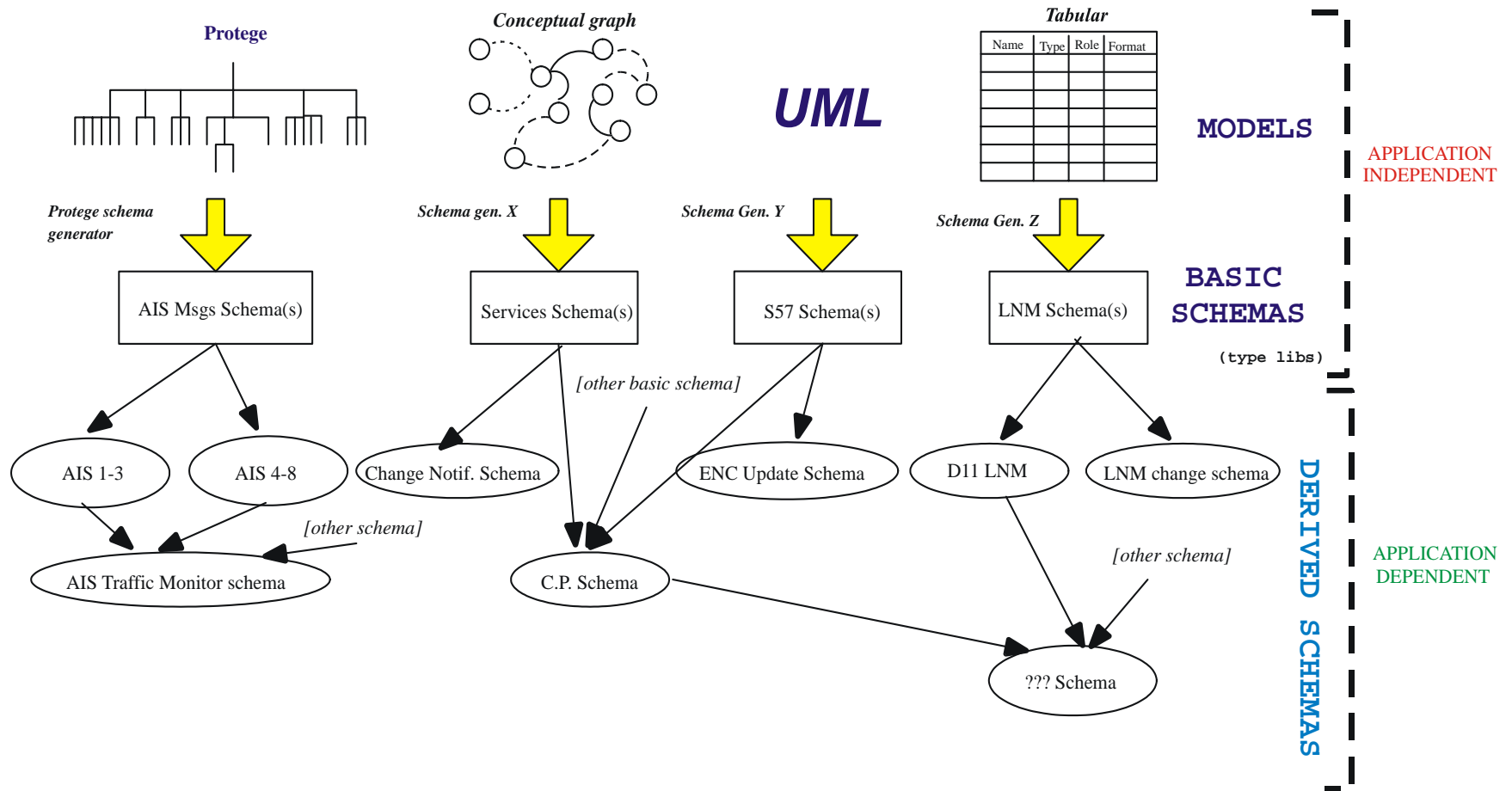
- 1) Namespace (new if first time)
- 2) XML “tagged” data
- 3) Name of Model or [Schema](#) and Namespace
- 4) Version Number
- 5) Date of Publication or revision
- 6) Owner Point of Contact
- 7) Intended Role or Scope

6.3 MIML Model Registry

The registry will be extracting “dictionary-type” information from the MIML Models and incorporating it in a registry database. This will enable searches for a particular class definition or such details. The registry will also provide a number of functions:

- 1) Search – All Users can search the WIN for models or schema.
- 2) Access – Most users can access basic MIML schemas.
- 3) Publish/Revise/Delete – A limited set of users can publish new models and basic schemas or revise/delete them.

Figure 2



6.3.1 Searching the MIML Registry

The searches should be permitted at various levels of the registry.

- a) Intra Model - i.e. searching for a particular class within the model. In this case the search result should return all information from all possible tables in that model related to that i.e. Class name, its parent class, documentation related to the class, slots in that class, instances, type....

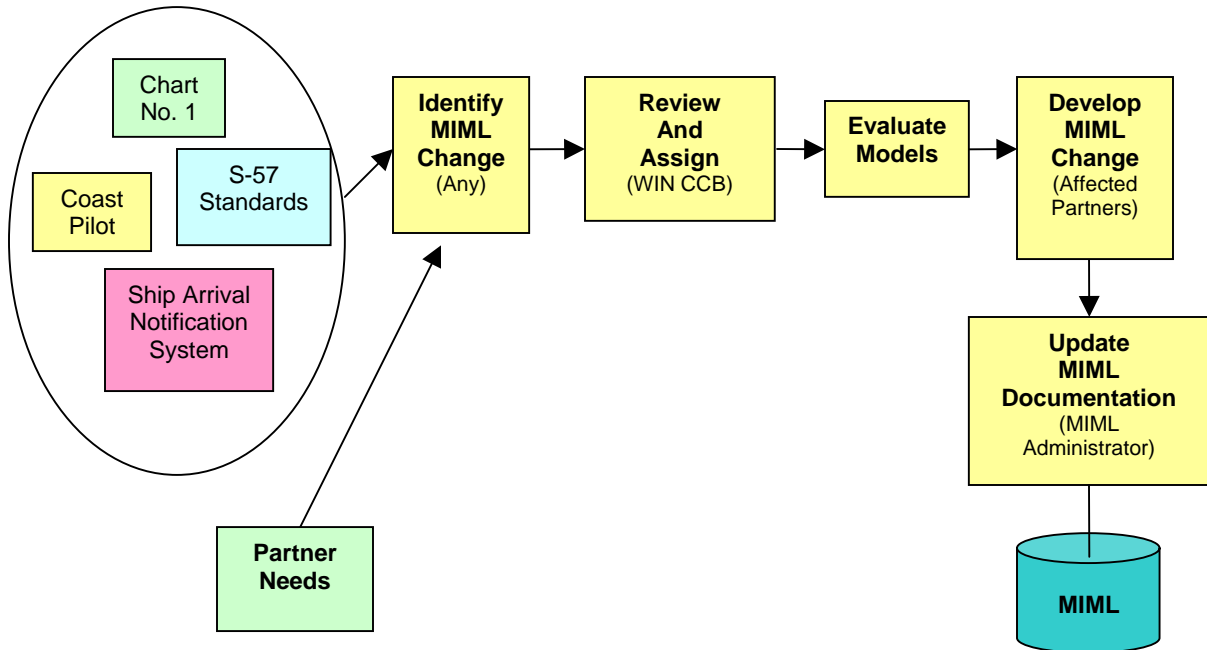
For intra model searching the class name can be preceded by the model name.

- b) Model search - i.e. It will search for a specific model, given the name. Search result will display all the classes and slots associated with that model.
- c) Inter Model - i.e. Searching for a particular class. All possible classes of the same name will be returned from all the models present in the registry

6.4 Maintenance of the MIML

Figure 3 shows an overview of the process. The expected steps in maintaining MIML are as follows:

Figure 3



6.4.1 Identify MIML Change

The WIN/MIML CCB reviews authoritative “Source Standards” for revisions that impact MIML:

- National Oceanic and Atmospheric Administration (NOAA), Office of Coast Survey (OCS), Coast Pilot_{tm}
- Chart No. 1, United States of America, Nautical Chart Symbols Abbreviations and Terms, Tenth Edition, November 1997, Prepared Jointly by Department of Commerce, National Oceanic and Atmospheric Administration, National Ocean Service and Department of Defense, National Imagery and Mapping Agency
- International Hydrographic Organization (IHO) Transfer Standard for Digital Hydrographic Data S-57 Edition 3.1 November 2000
- Ship Arrival Notification System (SANS)
- Local Notice to Mariner (LNM)

A WIN Partner initiates a request for correction, change or enhancement.

6.4.2 Review and Assign

The WIN/MIML CCB will:

DRAFT

- Review for applicability:
 - Source change
 - Change request
 - Evaluate impact and prioritize
 - Assign for action to affected partners
- 1.
 - 2.

6.4.3 Develop MIML Change

The MIML change potentially includes:

- Change to a MIML DTD or [Schema](#)
- Change to WIN Gateway to reflect the change
- Change to other affected site applications

6.5 Update MIML Documentation

The MIML Administrator will update the MIML Knowledge Base, including:

- Archive the current version
- Update [Model](#) as appropriate
- Modify affected DTD(s) or Schema
- Post to the WIN
- Notify Partners that the new version is available

For continuing operations, four primary functions will be performed to maintain the MIML:

- 1) Maintain Version Control
- 2) Track Requested Changes
- 3) Receive And Verify Changes
- 4) Incorporate changes and maintain the latest version of MIML on the WIN

6.6 MIML Assumptions

- 1) MIML will register models. These models will contain one or more object classes. We expect that many of these object classes will be duplications. These duplications will not be resolved when the models are registered to MIML. But

DRAFT

they may be resolved during subsequent model management.

- 2) MIML will support variable subsets of detail. The level of detail necessary for a schema will depend on the provider giving their customers the appropriate level of information.
- 3) One model can be used to create one or more schemas and a schema can be created from more than one model. A schema can be created from another schema. One schema can include another, and can use items from those schemas directly, or can extend or restrict element and attribute definitions from the included schema.

6.7 Current MIML

The MIML currently contains fourteen [models](#):

Model	Current Source	Future Source
AIS messages 1 –3	Protégé	NDRSMP
LNМ	Protégé	Various
SANS	Protégé	OSC
S57	Protégé	OpenEcdis
Coast Pilot	Protégé	Various
Chart 1	Protégé	NOAA
VTС	Protégé	PAWSA
Weather	Protégé	NWS
Marine Services	Protégé	Marine Exchanges
Documents for Mariners	Protégé	
Communications	Protégé	Deepwater
Port Information	Protégé	HMIS
Regulations	Protégé	M Portal
Vessel	MSMP	

The models are used to generate [schema](#), which describe the data available on WIN.

6.8 Future MIML

- 1) Individual VTS, Weather, Marine Services, Documents for Mariners Communications, Port Information and Regulations models will replace the Coast Pilot model.
- 2) The Chart 1 model will be replaced with the S57 model. New object classes will be added if necessary.
- 3) The OSC SANS model will replace the Protégé SANS model. A discussion will be held at OSC to discuss the Protégé SANS model and the MSMP Vessel model in relationship to the current OSC SANS model.

6.9 Outstanding MIML Issues

- 1) The European Union (EU) Maritime XML Project recommends for all geospatial information, to keep the models like S57. This is a structure of object classes and attributes. One table represents each object class, with the attributes being represented in the table columns. A new object is brought into the model by providing its class and attributes. In S57 there is a single record structure for transfer of S57 data. It mandates the organization of data in the record format and specifies mandatory object classes versus optional object classes.
- 2) A meeting of the WIN/MIML CCB has recommended allowing [owners](#) to provide data to the WIN in their own format (not necessarily XML). If this is allowed then we have the opportunity for translation organizations to exist as part of WIN.

6.10 WIN/MIML CCB Recommendations

- 1) Develop geolocation and time [models](#). These models will be available on WIN and they recommend the best practice for geolocation and time information.
- 2) MIML will use the XML representation of S57 produced by the Maritime XML project from the European OpenEcdis community.
- 3) Service Stores will be available on the WIN. There will always be a need to translate from one format to another. Organizations may be interested in researching other organizations data models. There will be an optional model organization on the WIN. An organization can choose to provide their model to the model organization.
- 4) MIML will recognize that conceptual models, facts, and relationships between objects may also be expressed in RDF (Resource Description Framework) format (www.w3c.org) or DAML (DARPA Agent Modeling Language) format (www.daml.org). Since these are even newer than XML we recognize that these might be used in the future.

7. MIML TECHNICAL APPROACH

7.1 Overall structure of technical solution

WIN will consist of an Internet-based server, which provides knowledge models and XML [schemas](#). The schemas are collectively called the Maritime Information Markup Language (MIML).

Users will be able to interface with the server via the Web interface and perform the following functions depending on their level of authority.

7.1.1 MIML Functions

- 1) It will provide a string based search mechanism by which one can look up a particular model.
- 2) If existing models in the registry fail to satisfy a users requirement, they can submit a new model to the registry. The different models will be given different namespaces thus avoiding any naming conflicts.
- 3) There will be a function as a format converter, verifier and parser.
- 4) There will be a function that parses the XML specification file and extracts the class names, attributes and other relevant information from this document and add them to the registry.

7.1.2 User Levels of Authority

The “users” of this system, fall in the following categories:

Level 1: [Model](#) and [schema](#) designers who can publish, revise and delete models and basic schemas.

Level 2: Application schema designers who can create application schemas from the models and basic schemas or published application schemas, but not add or revise models or basic schemas.

Level 3: Application developers who can search for models and schemas but not create or revise schemas or models.

Figure 2 shows the conceptual structure of the system. All the components shown in that figure reside on the Web server and can be accessed via the Web interface.

7.2 MIML Knowledge Models

Models can be in any “standard” form approved by the WIN/MIML CCB (see Section 1).

7.2.1 Basic Schemas

The “basic schemas” are type libraries that convert the models into XML-based descriptions of concepts in the model Domain (e.g., declare objects and attributes). The basic schemas are application-independent. The models and basic schemas describe the conceptual structure of their domains (e.g., weather, geospatial data, port and cargo services, etc.)

7.2.2 Derived Schemas

The derived schemas use one or more of the basic schemas, and are intended for structuring information and documents for direct use by applications, e.g., for data transfer. For example, the (hypothetical) “ENC update schema” in the figure would use the S57 schema in defining its own XML elements for updating ENCs (a sample element in this ENC update schema might describe an update to the status of a beacon as well as

the source of the update). The C.P. (Coast Pilot) schema in the figure would be derived from (use tags defined in) the S57 schema as well as the Services (port facilities) schema, because the Coast Pilot contains mentions of navigational objects (buoys, beacons, etc.) as well as descriptions of port facilities.

7.2.3 Schema Management

The WIN/MIML CCB will identify means of distinguishing models and schemas, specify the scope of individual models/schemas, and specify metadata and documentation that must be made available by model or schema “[owners](#)”.

The WIN/MIML CCB will develop other criteria for [models](#) and [schemas](#) (see Section 1). The definition of responsibilities for specific model and schema development, “[ownership](#)” and “[stakeholding](#)” in specific models, the approval process for models and schemas are WIN/MIML CCB responsibilities.

Arizona State University will provide functionality for generating basic schemas from Protege models or the Protege schema generator. We will be investigating convenient generators for other modeling tools (e.g., ORM, UML, ER) and incorporate them if available.

7.3 Dictionary/Registry of Knowledge Models

Protege's HTML generation function will produce the initial version of a dictionary or registry of model entities, attributes, and relationships. More documentation tools will be added. The owners are responsible for providing the model and schema metadata. See minimum model requirements (Section 1.2.1)

7.3.1 MIML Model Registry Submission Process

To submit models to the registry the following steps would need to be taken.

- 1) Register as a user.
- 2) Search through the registry to determine if the model already exists.
- 3) If the model doesn't exist, the user will submit it to the registry
- 4) The user will select a namespace or create a new namespace.
- 5) Build the submission package.
- 6) Put the package through the package verifier that will check for errors or inconsistencies in the package format.
- 7) After the package is successfully accepted into the registry, the sender will be notified about the progress.

Note: The submission package will be in XML format.

7.3.2 Preparation of submission package for the MIML registry

In order to submit a model to the MIML registry, the information has to be packaged in an XML format. For Protégé files, the information about the structure of the model will

be extracted from the .pont, .pins, and .prpj files using an appropriate standard XML notation to produce an XML schema file.

7.3.3 MIML Model Validation Rules

Following will be the validation checks performed on the files submitted to the registry.

- 1) Check for syntax. i.e. whether the tags are properly in place, missing tags etc.
- 2) Check for valid relations. The registry has a schema that will be static. All possible relations that can be expressed between the various tuples will be taken into consideration in the registry schema. So if any relation in the submitted document does not correspond to the standard relations defined, such errors should be corrected.

As of now the validation model is optional. It may not be included in the registry design. But the validations can be performed by the organizations wishing to submit their models to the registry by the WINCCB, before they are published. In addition, one can also use the IBM SQC or the XSV tool on the website at: <http://validator.w3.org/> for performing validations on the schemas.

7.3.4 Basic Structure of the Registry for the MIML Models

The registry will have a set of tables for each model. Each model may have many tables associated with it, depending on how the model is maintained: e.g. Class Table, Slot or Attribute Table, Meta Table. The following are examples based on the Protégé Model format:

- 1) Main Table:
Model Name, Version Number and the combination would be the primary key.
- 2) Meta Table:
Model Name and Version Number, Link to physical location of an HTML document generated by model repository that describes the classes, metaclasses, slots (attributes) and instances of the model. One could actually click on a class name to view its description etc. Also there is a link to the basic XML Schema file, a link to the physical location of the actual model in its native format (i.e. .pont, .prpj for project files) and the names of any other models. An example of the last would be the proposed geolocation model, which would be included in other models.
- 3) Class Table:
Class Name, Model Name, Default Class, Direct Parent Class, Role, Comment. The primary key would be the class and model name.
- 4) Slot Table:
Slot Name, Type, Cardinality, CreateAccessor, Allowed Classes, Class Name,

DRAFT

Model Name, Comment. The class and model name would be the primary key.

5) Instance Table:

Name, InstanceOf (slot/class), Classname, Slotname, ModelName, ClassInstanceValues

In the future we may re-design the registry tables using Object-Role Modeling (ORM - Dr. Terry Halpin, Information Modeling and Relational Databases - from Conceptual Analysis to Logical Design ISBN 1-55860-672-6 Academic Press 2001 copyright and Object Role Modeling, <http://www.orm.net>). This would remove the Slot Table because in this technique there are no attributes; everything is defined in terms of Objects and Roles.

7.3.5 Version Control

Schema and model editing will NOT be part of the initial functionality due to the difficulty of interfacing available tools (e.g., UML and conceptual graph editors with the Web backend). Instead, Level 1 and 2 users will need to create or edit models schemas outside of the Web interface and load them ("publish" them) after approval by the appropriate CCBs.

The WIN/MIML CCB will provide the ability to add the approved new models and schemas, or publish revised versions of models.

7.3.6 Model or Schema Derivation

The models or schema used and dependent on any model/schema will be visible by any user.

7.4 Model Redundancy

Ensuring that overlapping concepts are not defined in different models or schemas will be a major effort within WIN/MIML, although it can never be 100% effective.

Automated tools will be provided that can compare model structures, and entity/attribute definitions (documentation) to detect potential problems based on class name or other word comparisons.

7.5 Model Integration/Reconciliation

MIML models will be continually updated and new models will added. XML schemas will be generated from the models and used with the existing XML functionality of namespaces. The WIN/MIML CCB may need to devise and enforce criteria for namespaces and schema derivation (e.g., require version information as part of schema namespaces).

The peer-to-peer nature of the WIN will provide the opportunity for the user to select the

information from their preferred sources. If multiple vendors (data suppliers) come up with different [models](#) for the same domain, then either those models can be reconciled by the model [owners](#) themselves (outside of the Web server) to give a third, common model, or application designers/programmers can pick whichever they like and use the WIN translators to transform data from one format to another, or even use both models. All models will have different namespaces, so it will be possible to distinguish between data forms even if they cover the same domain.

New Owners (registers) to the WIN will have their own “namespace”. This allows for duplication of objects. Their own Configuration Control Board controls the developments of models of data provided to MIML.

8. MIML DEFINITIONS

Most of the definitions for the terms discussed in this document can be found in Glossaries of Reference 4 and 6. Following are some specific to the implementation of WIN and MIML.

Owner - provides data to WIN

Schema – a specific view of the model for use in particular computer programs.

Model - describes data provided to WIN

9. REFERENCES

1. DON Policy on the Use of Extensible Markup Language (XML) of 13 December 2002
2. Requirements for the DON XML Developer’s Guide, November 2002
3. Department of The Navy XML Developers Guide, Version 1.1, 1 May 02
4. Draft Federal XML Developer’s Guide of April 2002
5. Department of The Navy Vision for Extensible Markup Language (XML) 15 March 2002
6. Department of The Navy XML Developers Guide, 7 November, 2001
7. World Wide Web Consortium (W3C) Recommendation, XML Schema Part1: Structures, 2 May 01
8. W3C Recommendation, XML Schema Part 2: Datatypes, 2 May 01
9. W3C Recommendation, Extensible Markup Language XML 1.0 (Second Edition), 6 Oct 2000

Filename: MIMLv03.doc
Directory: C:\TEMP\IE6\OLK2BC
Template: U:\Templates\Normal.dot
Title: TABLE OF CONTENTS
Subject:
Author: Work Station 1
Keywords:
Comments:
Creation Date: 5/27/2003 11:07 AM
Change Number: 5
Last Saved On: 6/25/2003 9:50 AM
Last Saved By: KShea
Total Editing Time: 25 Minutes
Last Printed On: 6/25/2003 1:51 PM
As of Last Complete Printing
Number of Pages: 17
Number of Words: 4,373 (approx.)
Number of Characters: 24,929 (approx.)